Molecular cloning of complementary DNAs encoding two cationic peroxidases from cultivated peanut cells

(gene expression/ λ gt10/ λ gt11/secondary structure/Arachis hypogaea)

Dominique Buffard, Colette Breda, Robert B. van Huystee*, Omorefe Asemota[†], Michele Pierre, Duc B. Dang Ha, and Robert Esnault[‡]

Institut des Sciences Végétales, Centre National de la Recherche Scientifique, 91198 Gif-sur-Yvette Cedex, France

Communicated by J. E. Varner, August 17, 1990 (received for review April 25, 1990)

ABSTRACT We have isolated, cloned, and characterized two cDNAs corresponding to the mRNAs for cationic peroxidases synthesized by cultured peanut cells. The first clone was obtained from a phage λ gt11 library screened with antibodies directed against the major secreted isozyme. Its predicted amino acid sequence, deduced from the 1228-base-pair (bp) cDNA, revealed a 22-amino acid signal peptide and a 294amino acid mature protein $(M_r, 31, 228)$. The second clone was isolated from a λ gt10 library screened with oligonucleotides corresponding to the regions for acid/base catalysis and the fifth ligand of heme. This cDNA (1344 bp) encodes a protein (330 amino acids) with a mature peptide of 307 residues $(M_r,$ 32,954). The two peanut peroxidases are 46% homologous. The estimated gene copy numbers of these peroxidases might be close to 1 or 2 per haploid genome. A comparison of the amino acid sequence of these peanut peroxidases with other known isozymes shows two already known regions of homology (the region for acid/base catalysis and the fifth ligand of heme). Moreover, some new characteristics appeared such as a glycosylation site identical in five of the seven isozymes, a putative antigenic determinant common to all the isozymes, and a region of the highest homology. A secondary structure prediction showed that it corresponds to a 16-amino acid helix linked to the next one by a long stretch of β strands and coils and might represent a critical structural element.

Most higher plants possess a large number of peroxidase isozymes, which have been used as convenient enzyme markers in genetic, physiological, and pathological studies (1, 2). Their pattern of expression is influenced by environmental stimuli, is developmentally regulated, and is tissue specific (3), as exemplified in tomato, where anionic isozymes account for nearly 90% of the peroxidase activity in all tissues except for the root tissue (4).

The large number (up to 30) of isozymes could represent gene copy numbers through polymorphism and posttranslational modifications or artifacts of preparation such as crosslinking with phenols (2). These many possibilities of so-called isozyme expression make it difficult to study the actual function of peroxidase. Peroxidase has been implicated in polysaccharide cross-linkages with extensin monomers (3), indole-acetic acid oxidation, lignification, wound healing, phenol oxidation, defense against pathogens, and regulation of cell elongation (1).

Cloning of plant peroxidase could help in elucidating the physiological role that each of the isozymes plays in plant development, and the first report appeared on anionic peroxidase in tobacco (6). Insertion of this cDNA under control of cauliflower mosaic virus 35S promoter into the tobacco genome led to the expected overproduction of the isozyme but also to wilting at the time of flower bud initiation (7). These remarkable results illustrate how complex the analysis of plant peroxidase function is. Under normal conditions the anionic isozyme is associated with cell wall formation, particularly lignin synthesis (4), and is preferentially expressed in immature xylem and epidermis (7).

Cultured cells are able to respond rapidly to many environmental stimuli or to biological compounds such as elicitors (8). Therefore, they may serve as a simpler experimental model to analyze (i) the mechanisms by which plant cells perceive these signals and (ii) the series of cellular events that lead to modulation of transcription of specific genes. Peroxidases are not only synthesized by cultured cells but also are secreted into the medium (2, 9) from which they may be purified easily. Extracellular peroxidase activity has a strong correlation to growth kinetics and cell dry weight (10).

Peanut cell peroxidases have been studied in detail (2), and it was shown that the released isozymes consisted of two cationic and one anionic species. The major cationic isozyme represents 75% of the medium peroxidase activity. Up to 2% of the total $poly(A)^+$ mRNAs isolated from the cultured cells are engaged in its synthesis (11). These results raised some intriguing questions. Is the extracellular location of the cationic isozyme related to a specific function for the cultured cells? Are its high rate of synthesis and secretion directly related to the mode of growth of these cells? Is there only one major cationic peroxidase synthesized by these cells and, if so, what are the mechanisms of activation and repression of the various peroxidase genes? Answering these questions would broaden our knowledge of the peroxidase function. To this end we undertook the cloning of the major secreted cationic isozyme from peanut cells but also extended our search to other genes to analyze their differential expression. In this communication we present our first set of results devoted to two different cDNAs encoding cationic peroxidases, one of them being the major secreted isozyme.[§]

MATERIALS AND METHODS

Plant Cells. Culture of peanut cells, Arachis hypogaea, was as described (12). Cells were collected (after 6 or 7 days of subculture) by filtration through a sterile 3MM Whatman disk and frozen at -180° C.

Preparation of RNA. Frozen cells (10 g fresh weight) were grown with a Polytron in 25 ml of 6 M guanidinium chloride/ 0.1 M potassium acetate. Total RNA was extracted by the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. \$1734 solely to indicate this fact.

^{*}Present address: Department of Plant Sciences, University of Western Ontario, London, ON, Canada.

[†]Present address: Physiology Division, National Institute for Oil-Palm Research, PMB 1030, Benin City, Nigeria.

[‡]To whom reprint requests should be addressed.

[§]The sequences reported in this paper have been deposited in the EMBL/GenBank data base (accession nos. M37636 for PNC1 and M37637 for PNC2).

guanidinium salt procedure (13). The $poly(A)^+$ RNA was selected by oligo(dT)-cellulose chromatography.

Construction of the cDNA Libraries. Double-stranded cDNA of $poly(A)^+$ RNA was synthesized by using the Amersham cDNA synthesis system and was used to prepare phage $\lambda gt10$ and $\lambda gt11$ libraries (Amersham systems).

Antibody Selection of λ gt11 Recombinants. The λ gt11 clones were plated on *Escherichia coli* strain 1090 to yield 3 \times 10⁴ plaque-forming units per 100-mm plate and were screened as described by Huynh *et al.* (14). Immunodetection was carried out as described (15).

Oligonucleotide Screening. Mixed oligonucleotides, synthesized by Appligène (Strasbourg, France) were end-labeled with T4 polynucleotide kinase and $[\gamma^{-32}P]ATP$ at 3000 Ci/ mmol (1 Ci = 37 GBq) from NEN. The ampicillin-resistant transformants were screened by plaque hybridization with the oligonucleotides mixtures as a probe by the method of Woods *et al.* (16) with hybridization temperatures of 37°C and 42°C for the 17- and 23-mers, respectively. Final washings were done at 47°C and 52°C, respectively.

Subcloning and DNA Sequencing. After the subcloning into M13mp19, deletions of the inserts were prepared using the Cyclone System from IBI. Nucleotide sequences were determined by the dideoxynucleotide chain-termination method {T7 sequencing kit from Pharmacia and deoxy-adenosine 5'-[α -(35 S)thio]triphosphate at 1500 Ci/mmol from NEN}.

DNA Filter Hybridization. Peanut cell high molecular weight DNA was isolated as described by Richards (17). DNA (10 μ g) was cleaved with various restriction enzymes and subjected to 0.6% agarose gel electrophoresis. The DNA was then transferred to Hybond N (Amersham) filters and hybridized with cDNA inserts (labeled with the Amersham multiprime DNA labeling system and [α -³²P]dCTP at 800 Ci/mmol from NEN) at a specific activity of 1.8 × 10⁸ cpm/ μ g (prxPNC1 clone) or 1.8 × 10⁹ cpm/ μ g (prxPNC2) at a temperature of 42°C in 50% formamide according to Amersham's instructions. For washing, high-stringency conditions were used—i.e., the final wash was at 62°C in 0.018 M NaCl/1 mM phosphate, pH 7.4/0.1 mM EDTA/0.1% sodium dodecyl sulfate (SDS). A quantitative analysis was done on the genomic blots with a BI0COM 200 apparatus.

Computer Analysis. Analyses were made by using the PCGENE software (Genofit) or the methods developed by Biou *et al.* (18) for secondary structure prediction.

RESULTS

Cloning Strategy. At the outset of this work all of the available molecular data consisted of the amino acid sequences from the horseradish (19) and turnip (20) peroxidases. Thus, our strategy was based on the preparation of a $\lambda gt11$ library, the screening of which was to be conducted with polyclonal antibodies directed against the glycosylated or chemically deglycosylated (21) form of the major cationic peroxidase secreted by the cultured peanut cells.

A significant comparison of some regions that might have been conserved (i.e., the active site) has been made possible through the publication of the cDNA sequences encoding two anionic peroxidases—one from tobacco (6) and the other from potato (22). The deduced oligonucleotide probes corresponding to the regions for acid/base catalysis and the fifth ligand of heme and designated as "acid/base probe" and "heme fixation probe" (Fig. 1) were used to check, first of all, the tentatively positive inserts detected from the $\lambda gt11$ bank. Moreover, they made possible the use of a $\lambda gt10$ library to search for related core sequences encoding polypeptides that might not respond to the available antibodies.

Screening of the λ gt11 Library. A cDNA library consisting of 5×10^6 plaque-forming units/µg of λ DNA was constructed in λ gt11. One-fifth of the library was plated; after two rounds of screening with antibodies directed against the native per-

Γ	Horseradish	PHE	HIS	-ASP	-CYS	-PHE	-VAL	VAL-ALA-LEU-SER-GLY-GLY-HIS-THR											
	Turnip	-	-	•	-	-	-			•	•	•	-	-	ALA	•	-		
1	Tobacco	.	-	-	-	•	-			•	•	-	-	-	ALA	•	-		
	Potato	-	-	-	-	-	•			•	-	•	ALA	•	ALA	•	-		
	Tobacco	TTC	CAT	GAT	TGT	TTT	GTT	Γ		GTT	GCT	CTA	TCA	GGT	GCA	CAC	ACA		
B	Potato			c	c	c	•••				G		GCT		G		G		
Γ	Synthesized					TTC		(8)	((GTT 	GCA T	СТА Т	TCA T	GGA T	GCA T	CAC T	AC 		
C	Oligonucleotides	1	7 ME	RS F ase	ROBE	:		(Ь)	{ {		2	23 M	ERS 1	PROBI	GCG T S Probe		AC 		

FIG. 1. DNA sequence of the oligonucleotide probes for screening peanut cDNA clones. (A) Direct (horseradish and turnip) or deduced (tobacco and potato) amino acid sequences of the regions corresponding to acid/base catalysis and the fifth ligand of heme. (B) The nucleotide sequence of the corresponding tobacco and potato cDNAs. (C) The oligonucleotide probes: the acid/base probe is a mixture of 32 combinations of 17-mers; the heme fixation probe consists of two sets (a and b) of 64 combinations of 23-mers synthesized independently but mixed together for the screening.

oxidase, 20 positive clones were identified with inserts ranging from 0.7 to 2 kilobases (kb). As the deglycosylated protein has a M_r of 31,500 (21), nine phages bearing an insert from 0.9 to 1.2 kb were submitted to a new round of immunodetection, this time using the antibodies directed against the deglycosylated peroxidase. Seven of them gave a positive response, and the phage DNAs from four plaque-purified individual clones were isolated for subcloning and DNA sequencing.

Nucleotide Sequence Analysis of the prxPNC1 Clone. A strong cross-hybridization was observed between the four selected inserts, and partial sequencing of these four clones indicated that they were derived from the same RNA. The complete nucleotide sequence of the longest cDNA insert [1128 base pairs (bp)], noted as prxPNC1, and the deduced amino acid sequence of this clone are shown in Fig. 2. Hybridization with the oligonucleotide probes (not shown) and comparison with the previously published sequences (6, 19, 22, 23) indicated that this is indeed a peroxidase cDNA.

The predicted size of the mature peroxidase is 294 amino acids with a M_r of 31,228. It is predicted from the cDNA sequence that the protein is synthesized as a preprotein of M_r 33,549 with a predominantly hydrophobic 22 amino acid signal sequence (Fig. 2). The mature protein, of cationic nature (theoritical pI, 7.8), is secreted as a glycoprotein with four N-glycosylation sites (at residues 60, 142, 185, and 275), as also found by protein hydrolysis (21).

Screening of the λ gt10 Library and Nucleotide Sequence Analysis of the prxPNC2 Clone. The λ gt10 library consists of 4.5×10^5 plaque-forming units/µg of λ DNA, and probing was done on half of the library with the "acid/base probe." Seven positive plaques remained after a second screening, corresponding to phages bearing an insert from 0.8 to 2 kb. We selected one plaque-purified phage bearing an insert of around 1.3 kb, designated prxPNC2, that does not crosshybridize with prxPNC1 in stringent conditions (not shown). The nucleotide and deduced amino acid sequences of the prxPNC2 clone are presented in Fig. 3. The cDNA encodes a preprotein of M_r 35,556 (330 amino acids) with a putative signal sequence of 23 residues (Fig. 3). The mature polypeptide (307 amino acids, M_r 32,954) is also of cationic nature (theoretical pI, 8.5), and it bears only one glycosylation site (at position 189).

The cDNA clone is 1344 bp long and possesses a 199-bp sequence upstream of the starting ATG, which is characterized by a long repeated sequence of 69 nucleotides (boxed in Fig. 3) rich in A+T (73.9%). The 155-bp 3' untranslated

			-	22																					-1	+1																
[1]	TCAT	AGCC/	AGC	ATG	GCA	CTT	CCA	ATT	AGC		GTT	GAT	TTC	TTA	ATA	TTT	ATG	TGT	CTT	ATA	GGA	TTA	GGG	TCA	GCT	CAA	TTG	TCA	TCT	AAT	TTT	TAT	GCC	ACA	***	TGT	ccc	AAT	GCA	CTT	TCA	[126]
				Net	Ala	Leu	Pro	Ile	Ser	Lys	: Val	Asp	Phe	Leu	Ile	Phe	Met	Cys	Leu	Ile	Gly	Leu	Gly	Ser	Ala	Gln	Leu	Ser	Ser	Asn	Phe	Туг	Ala	Thr	Lys	Cys	Рго	Asn	Ala	Leu	Ser	
				20																				40											•	•						
[127]	ACA	ATT /	MG	TCA	GCA	GTG	AAC	TCC	TGT	GTC	GCC		GAA	GCT	CGC	ATG	GGA	GCT	TCC	CTT	CTT	CGC	CTT	CAT	TTT	CAT	GAT	TGC	TTT	GTT	CAA	GGA	TGT	GAT	GCA	TCA	GTG	CTA	TTA	GAT	GAT	[249]
	Thr	Ile l	.ys	Ser	Ala	Val	Asn	Ser	Cys	: Val	Ala	Lys	Glu	Ala	Arg	Net	Gly	Ala	Ser	Leu	Leu	Arg	Leu	His	Phe	His	Asp	Cys	Phe	Val	Gln	Gly	Cys	Asp	Ala	Ser	Val	Leu	Leu	Asp	Asp	
			60																				80																			
[250]																																										[372]
	Thr	Ser /	۱sn	Phe	Thr	Gly	Glu	Lys	Thr	Ala	ı Gly	Pro	Asn	Ala	Asn	Ser	Ile	Arg	Gly	Phe	Glu	Val	Ile	Asp	Thr	Ile	Lys	Ser	Gln	Val	Glu	Ser	Leu	Cys	Рго	Gly	Val	Val	Ser	Cys	Ala	
		100																				120																				
[373]	GAT	ATT (CTT	GCT	GTT	GCT	GCT	AGA	GAC	: TCT	GTT	GTT	GCT	CTA	GGA	GGA	GCA	AGT	TGG	AAT	GTG	TTA	TTG	GGA	AGA	AGA	GAC	TCA	ACC	ACT	GCA	AGT	TTA	AGC	TCT	GCT	AAC	TCA	GAT	TTG	CCG	[495]
	Asp	Ile l	.eu	Ala	Val	Ala	Ala	Arg	Asp	Ser	· Val	Val	Ala	Leu	Gly	Gly	Ala	Ser	Тгр	Asn	Val	Leu	Leu	Gly	Arg	Arg	Asp	Ser	Thr	Thr	Ala	Ser	Leu	Ser	Ser	Ala	Asn	Ser	Asp	Leu	Pro	
	140																					160		* *	* *	• •	• •	• •	• •	* *	٠											
[496]																																										[618]
		Pro F	he	Phe	Asn	Leu	Ser	Gly	Leu	ı Ile	e Ser	Ala	Phe	Ser	Asn	Lys	Gly	Phe	Thr	Thr	Lys	Glu	Leu	Val	Thr	Leu	Ser	Gly	Ala	Ħis	Thr	Ile	Gly	Gln	Ala	Gln	Cys	Thr	Ala	Phe	Arg	
	180																				200																				220	
[619]																																										[741]
	Thr	Arg 1	lle	Tyr	Asn	Glu	Ser	Asn	lle	: Asp	Pro	Thr	Tyr	Ala	Lys	Ser	Leu	Gln	Ala	Asn	Cys	Рго	Ser	Val	Gly	Gly	Asp	Tinr	Asn	Leu	Ser	Pro	Phe	Asp	Val	Thr	Thr	Рго	Asn	Lys	Phe	
																				240																					260	
[742]																																										[864]
	Asp	Asn /	lla	Tyr	Tyr	Ile	Asn	Leu	I Arg	Asr	Lys	Lys	Gly	Leu	Leu	His	Ser	Asp	Gln	Gln	Leu	Phe	Asn	Gly	Val	Ser	Thr	Asp	Ser	Gln	Val	Thr	Ala	Tyr	Ser	Asn	Asn	Ala	Ala	Thr '	Phe	
																				280																						
[865]																																				~~~		ATT	GTGG	JATTTA:	TAT	[995]
		Thr /																																								
[996]		CATG	IIGG	GIA	AAA		AGTT	GTGG	GAAT	TTT	ATGT	ATTI	CTAG	ATGT/	ATT	CAGTI	ACA	TGT/	IGAA	SC/A	ATTG	TTTC	CTCA	TCAA	TTGT	AATG	CAAA	TTTC	AGTT	TCAA	GTAN	ww	ww	N []	128]							

FIG. 2. Nucleotide and deduced amino acid sequence of the prxPNC1 cDNA clone. The nucleotide sequence is presented over the amino acid sequence for the precursor protein, the start of the mature protein being at position +1. The localization of the oligonucleotide probes (see Fig. 1) is indicated by \blacksquare for the acid/base probe and \blacklozenge for the heme fixation probe. A putative polyadenylylation signal is underlined.

region contains a consensus polyadenylylation signal (AATAAA) 27 nucleotides before the addition of poly(A).

Comparison of the Peroxidases. A comparison, based on the matrix used by Risler *et al.* (24), of the mature peanut cationic peroxidases, designated PNC1 and PNC2, respectively, with already published complete sequences (6, 19, 20, 22, 23, 25) is shown in Fig. 4, the partial sequence of wheat peroxidase (26) also being included. The PNC1 was found to be 46% homologous to PNC2, 61% to turnip TP7, and \approx 50% to the others. The PNC2 is significantly less homologous to other peroxidases, being from 33% (tomato or potato isozymes) to 47% (turnip TP7) homologous.

All of the peroxidases have eight cysteine residues, which are located in similar positions in the primary sequences and two invariable histidine residues (at positions 42 and 169 in PCN1, Fig. 4) that have been inferred in the active-site structure.

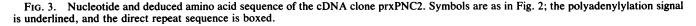
Three rather large domains of homology can be noticed. In the first one (residues 33–56), the peroxidases are 62–83% homologous; the histidine involved in the acid/base catalysis belongs to a subdomain of 100% homology: Phe-His-Asp-Cys-Phe-Val. In the second region (residues 91–110), the homology ranges from 61% to 95% with a sequence of seven residues (Val-Ser-Cys-Ala-Asp-Ile-Leu) present in six of the seven peroxidases and in the wheat sequence; the only discrepancy is that of horseradish peroxidase, in which there is a replacement of the isoleucine by a leucine. The third domain (12 residues from 161 to 172) includes the histidine residue involved in the fifth ligand of heme and shows more variability, with homology ranging from 58% to 75%. Among the shortest regions (121–127 and 138–142), the first one (Leu-Gly-Arg-Arg-Asp-Ser in PNC1) might be an important structural element, as it corresponds to a tentative antigenic determinant, the location of which is common to all of the analyzed peroxidases.

The number of glycosylation sites varies greatly according to the isozyme (from one in PNC2 and turnip to eight in horseradish). However, a glycosylation site (Asp-185) is common to five of the seven peroxidases.

A separate secondary structure prediction on PNC1 and PNC2, based on the methods of Biou *et al.* (18), is shown in Fig. 5. A striking similarity appears in the relative distribution of predicted helices in both sequences, despite the uncertainties linked to the secondary structure predictions. These predicted secondary structures are in good agreement with the tertiary structure model of Welinder (27). For example, the region of strongest homology (91 to 106 in Fig. 4) corresponds to an helix of 16 residues (helix D) linked to the next one (helix E) by a stretch of around 30 residues consisting of β -strands and coils. This might represent a critical structural element.

Estimation of Gene Copy Numbers. Restriction digests of peanut DNA and reconstruction standards were hybridized with the two peroxidase cDNAs prxPNC1 and prxPNC2. As shown in Fig. 6, a different set of fragments is hybridizable to each of the two probes, confirming the lack of cross-hybridization between the two clones in our conditions. A densitometric comparison of the hybridization signals suggests that the copy numbers per hyploid genome is around 1 for prxPNC2 and 2–3 for prxPNC1.

[1]	ATATCACATACAT	TTTCTATCTCA	TTCATCCTTA	ACATACAA	TATAAACA	ACTITAAT	TTCTCAA	TCATCA	TCATCA	TACT	TTCTA	TCTCAT	TCATCC	TTAACA	TACAAT	TAAAC	ACTT	AATTT	CTCAAC	TCATC	ATCATO	ATACT	AGCT	AGCTATT	T [163]
[164]	GTGCTTGATATTCA	ATTCTTAGTTAG	CAAAATAACA	-23 ATG GAG Met Glu																					
[298]	GTA GGG TTC TA Val Gly Phe Ty																								
[424]	GTC CAA GGC TG Val Gin Giy Cy																								
(550)	CCC GGT GTT GT Pro Gly Val Va																								
[676]	GTA AGC AAC TT Val Ser Asn Le 180																								6
[802]	CAA TTC TTC AG Gln Phe Phe Se																								
(928)	GCA CTC GAT AC Ala Leu Asp Th																								
[1054]	AGA TAT TTG GO Arg Tyr Leu Gl																								
[1180]	GCT TTT AAC TA Ala Phe Asn Er		IGTTTAAGTGT	ATTCATAG	GTATATTT	GTTAGTCT	TATTGAT	*****	TCTTTG	TTATTO	STITCI	CTATA	GAGAAT	CAATTT	CCAAGT	GTAATA	TTCC	ATAAA	TAATAT	CACAT	GTTTTI	TATTA		*****	A [1344]



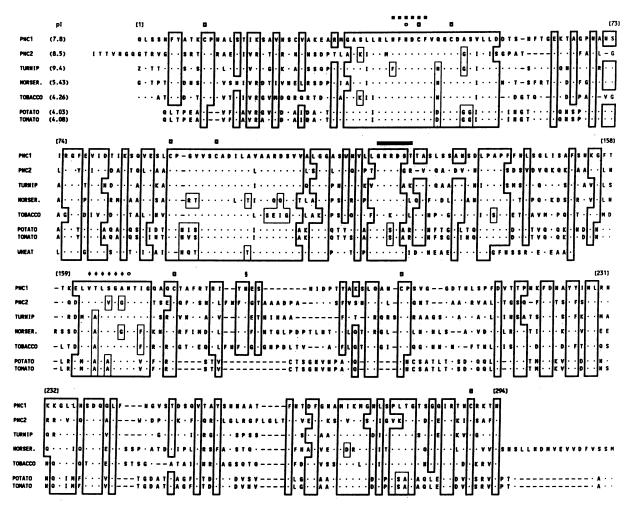


FIG. 4. Amino acid homology between seven mature and complete plant peroxidases. Alignment of the two peanut peroxidases (PNC1 and PNC2) is with the turnip [TP7] (20), tobacco (6), horseradish [C1] (23), potato [AP] (22), and tomato [Tap1] (25) isozymes; the partial wheat sequence (26) is also included for comparison. The calculated pl is indicated for each isozyme. The identical amino acids are designated by a dot, and gaps (introduced to produce the best alignment) are indicated by dashes. The position of the two used oligonucleotide probes is shown by \blacksquare and \blacklozenge (see Figs. 2 and 3). Regions boxed (heavy lines) indicate homology based on equivalent amino acids [F/Y, K/Q/R, S/T, V/I/L/(M)] as described by Risler *et al.* (24); nonequivalent amino acids in these regions are boxed by lighter lines. Conserved cysteines and histidines are noted by \blacksquare and \circ , respectively. A putative antigenic determinant present in all sequences is noted by a heavy bar (\blacksquare) and a common glycosylation site by §.

DISCUSSION

By using antibodies directed against the major isozyme secreted by peanut cells and oligonucleotide probes (con-

structed from the comparison with available peroxidase sequences) for the screening of $\lambda gt11$ and $\lambda gt10$ libraries, respectively, we have isolated several clones, two of which were thoroughly analyzed. The first clone (prxPNC1) was

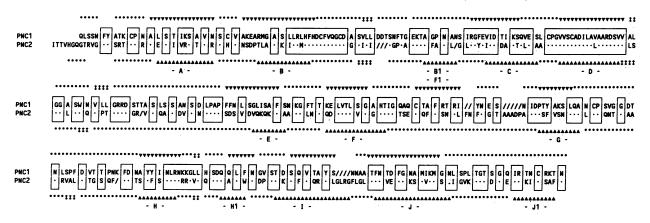


FIG. 5. Alignment of the two peanut peroxidases and predicted secondary structure. Identical amino acids are indicated by a dot, and a gap is indicated by a slash. The homologous regions (defined as in Fig. 4) are boxed. The first and fourth lines show the predicted secondary structure of the PNC1 and PNC2 peroxidases, respectively. Helices are represented by \checkmark for PNC1 and \blacktriangle for PNC2, coils by \circ , and beta strands by \updownarrow . The numbering of the helices is according to Welinder (27).

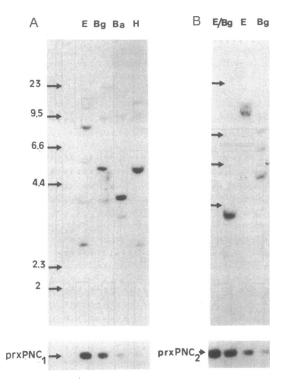


FIG. 6. Southern blot analysis of peanut DNA with peroxidase cDNA clones. Samples (10 μ g) of DNA from peanut cells were digested with restriction endonuclease enzymes singly or in combination prior to agarose electrophoresis, blotting, and hybridization with labeled cDNA probes prxPNC1 (A) and prxPNC2 (B). For copy number standards, purified *Eco*RI inserts [equivalent to 0.5, 1, 2, and 4 (prxPNC1) or 0.175, 0.35, 0.7, and 1.4 (prxPNC2) copies per haploid genome] were loaded onto the gel and were hybridized to the corresponding cDNA. Sizes are shown in kbp. Ba, *Bam*HI; Bg, *Bgl* II; E, *Eco*RI; H, *Hind*III.

detected in both libraries and the second (prxPNC2) was isolated from the λ gt10 library. They were sequenced in their entirety.

Homology of the deduced amino acid sequences between the two peanut peroxidases, which are synthesized as preproteins with a signal peptide of around 22 residues, is about 44%. The overall homology among the peanut isozymes and the other five known sequences (Fig. 4) ranges from 33% (PNC2 vs. highly anionic peroxidases from potato or tomato) to 61% (PNC1 vs. turnip TP7). However, amino acid alignment around the functional histidine residues (19)—i.e., His-42 and His-168, was conserved in all of the sequences and shows three rather large regions of homology.

The fifth ligand of heme region (residues 161-172) is much more variable than the others, as shown in Fig. 4, and that is confirmed by the published partial sequence of cucumber (5).

Two sets of six or seven amino acids are very highly conserved. The first one (Phe-His-Asp-Cys-Phe-Val) belongs to the acid/base catalysis region (residues 33–56) and includes the functional histidine residue. The second (Val-Ser-Cys-Ala-Asp-Ile-Leu), belonging to the third region of homology (residues 91–110), exists in seven of the eight isozymes, the horseradish C1 having a leucine (Fig. 4) and the C2 and C3 isozymes (not shown) having a valine instead of the isoleucine (23). According to secondary structure analysis (Fig. 5), this region consists of a helix (D) followed by a long stretch of β -strands and coils linking it to the next helix (E). This in excellent agreement with the peroxidase tertiary structure model built by Welinder (27) from a comparison of the horseradish and turnip proteins with the yeast cytochrome c peroxidase, which suggests that helix D is deeply buried and that its conserved sequence might indicate essential peptide-peptide contacts. The other two common characteristics noted in Fig. 4 (putative antigenic determinant and glycosylation site) are easily located on the surface of the structural model and correspond to exposed regions. This glycosylation site starting with Asn-185 [Asn-Xaa-(Ser or Thr)] is very close to Arg-182 and Tyr-184 that are putative binding sites of the substrate (5) and are conserved in all the sequences except the highly anionic isozymes from potato and tomato (Fig. 4).

Genomic southern mapping experiments showed that the gene copy numbers might correspond to 1 or 2–3. These values are in reasonable agreement with those available for tobacco [two copies of two parental peroxidase genes (6)] or tomato [two copies each of highly homologous genes (23)].

Unlike the two anionic isozymes from tomato (25) and the three moderately anionic ones from horseradish (23), which correspond in each case to very highly homologous polypeptides, the two peanut cDNA clones are encoding distinct molecular species of cationic nature, one of them being probably the major secreted peroxidase.

We are indebted to the Ministère des Affaires Etrangères (Paris) for the grant allocated to O.A., the Ministère de l'Education Nationale (Paris) for the invitation as a Professor to R.B.v.H., and the University of Paris 7 for support to D.B. and R.E.

- 1. Greppin, H., Penel, G. & Gaspar, T., eds. (1986) Molecular and Physiological Aspects of Plant Peroxidases (Univ. of Geneva Press, Geneva).
- ² 2. van Huystee, R. B. (1987) Annu. Rev. Plant Physiol. 38, 205-217.
- Cassab, G. I. & Varner, J. E. (1988) Annu. Rev. Plant Physiol. 39, 321-353.
- 4. Lagrimini, L. M. & Rothstein, S. (1987) Plant Physiol. 84, 438-442.
- Abeles, F. B., Dunn, L. J., Morgens, P., Callahan, A., Dinterman, R. E. & Schmidt, J. (1988) Plant Physiol. 87, 609-615.
- Lagrimini, L. M., Burkhart, W., Moyer, M. & Rothstein, S. (1987) Proc. Natl. Acad. Sci. USA 84, 7542–7546.
- 7. Lagrimini, L. M., Bradford, S. & Rothstein, S. (1990) Plant Cell 2, 7-18.
- 8. Templeton, M. D. & Lamb, C. J. (1988) Plant Cell Environ. 11, 395-401.
- 9. Mäder, M. & Walter, C. (1986) Planta 169, 273-277
- Shetty, K., Bothra, D., Crawford, D. L. & Korus, R. A. (1990) Appl. Biochem. Biotech. 24/25, 213-221.
- 11. Stephan, D. & van Huystee, R. B. (1981) Z. Pflanzenphysiol. 101, 313-321.
- 12. Kossatz, V. C. & van Huystee, R. B. (1976) Can. J. Bot. 54, 2089-2094.
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) Biochemistry 18, 5294-5299.
- Huyhn, T. V., Young, R. A. & Davies, R. W. (1985) in DNA Cloning: A Practical Approach, ed. Glover, D. M. (IRL, Oxford), Vol. 1, pp. 49-78.
- van Huystee, R. B., Breda, C., Sesto, P., Beopoulos, N. & Esnault, R. (1990) Plant Sci. 69, 19-26.
- Woods, D. E., Markham, A. F., Ricker, A. T., Goldberger, G. & Colten, H. R. (1982) Proc. Natl. Acad. Sci. USA 79, 5661–5665.
- Richards, E. (1987) in Current Protocols in Molecular Biology, eds. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K. (Wiley, New York), Vol. 1, pp. 2.3.1.– 2.3.3.
- Biou, V., Gibrat, J. F., Levin, J. M., Robson, M. & Garnier, J. (1988) Protein Engin. 2, 185-191.
- 19. Welinder, K. G. (1976) FEBS Lett. 72, 19-23.
- 20. Mazza, G. & Welinder, K. G. (1980) Eur. J. Biochem. 108, 481-489.
- van Huystee, R. B., Hu, C. & Sesto, P. A. (1990) in *Isozymes: Current Topics in Biological and Medical Research*, eds. Ratazzi, M. C., Scandalios, J. G. & Whitt, G. S. (Liss, New York), pp. 315-325.
- 22. Roberts, E., Kutchan, T. & Kolattukudy, P. E. (1988) Plant Mol. Biol. 11, 15-26.
- Fujiyama, K., Takemura, H., Shibayama, S., Kobayashi, K., Choi, J. K., Shinmyo, A., Takano, M., Yamada, Y. & Okada, M. (1988) Eur. J. Biochem. 73, 681-687.
- Risler, J. L., Delorme, M. O., Delacroix, H. & Hénaut, A. (1988) J. Mol. Biol. 204, 1019–1029.
- 25. Roberts, E. & Kolattukudy, P. E. (1989) Mol. Gen. Genet. 217, 223-232.
- Schweizer, P., Hunziker, W. & Mösinger, E. (1989) Plant Mol. Biol. 12, 643-654.
- 27. Welinder, K. G. (1986) Eur. J. Biochem. 151, 497-504.